What is claimed is:

1. A method of processing information related to documents in a collection of linked documents, the method comprising:

accessing a link log, the link log comprising a plurality of link records, wherein each link record comprises a respective source document identifier corresponding to a respective source document address and a respective list of target document identifiers corresponding to respective target document addresses; and

outputting a sorted anchor map, wherein the sorted anchor map comprises a plurality of anchor records, each anchor record comprising a respective target document identifier corresponding to a respective target document address and a respective list of source document identifiers corresponding to a respective list of source document addresses;

wherein the plurality of anchor records are ordered in the sorted anchor map based, at least in part, on their respective target document identifiers;

wherein, for at least one anchor record, a document located at the source document address corresponding to a source document identifier in the list of source document identifiers contains at least one outbound link, the at least one outbound link pointing to a corresponding target document address, the target document address corresponding to the respective target document identifier for the at least one anchor record.

2. The method of claim 1,

wherein each anchor record in the sorted anchor map further comprises a respective list of annotations.

3. The method of claim 2,

wherein the at least one outbound link is annotated by an annotation, the annotation included in the respective list of annotations for the at least one anchor record.

4. The method of claim 2, wherein at least one entry in the respective list of annotations for an anchor record in the sorted anchor map includes a text passage and a list of attributes of the text passage.

5. The method of claim 4, wherein the text passage is determined from text within a predetermined distance of an anchor tag in a source document in the collection of documents.

6. The method of claim 1, further including repeating the accessing and outputting so as to produce a layered set of sorted anchor maps.

7. The method of claim 6, further including,

when a merge condition has been satisfied, merging a subset of the layered set of sorted anchor maps, producing a merged anchor map;

wherein the merged anchor map includes a plurality of merged anchor map records, each merged anchor record corresponding to at least one anchor record from the subset of the layered set of sorted anchor maps, wherein the merged anchor records are ordered in the merged anchor map based on respective target document identifiers.

8. The method of claim 1, further including outputting a sorted link map, the sorted link map comprising a plurality of link map records, each link map record comprising a respective source document identifier and a respective list of target document identifiers.

9. The method of claim 8, further including repeating the accessing, outputting a sorted anchor map, and outputting a sorted link map so as to produce a layered set of sorted anchor maps and a layered set of sorted link maps.

10. The method of claim 9, further including

when a merge condition has been satisfied, merging a subset of the layered set of sorted link maps, producing a merged link map;

wherein the merged link map includes a plurality of merged link map records, each merged link record corresponding to at least one link record from the subset of the layered set of sorted link maps, wherein the merged link records are ordered in the merged link map based on respective source document identifiers.

11. The method of claim 10, wherein merging a subset of the sorted link maps further includes:

searching, within the subset of sorted link maps, for link map records containing a particular source document identifier; and

when a first link map record and a second link map record each contain the particular source document identifier,

if a target document identifier is contained in the list of target document identifiers in the first link map record and the target document identifier is not contained in the list of target document identifiers in the second link map record,

generating a delete entry in a record, the record comprising the particular source document identifier and the target document identifier.

12. The method of claim 11, further including

when an anchor merge condition has been satisfied, merging a subset of the layered set of sorted anchor maps, producing a merged anchor map;

where the merged anchor map includes a plurality of merged anchor map records, each merged anchor map record corresponding to at least one anchor map record from the subset of the layered set of sorted anchor maps, wherein the merged anchor records are ordered in the merged anchor map based on respective target document identifiers.

13. The method of claim 12, wherein, if a delete entry has been generated, a merged anchor map record containing the target document identifier in the delete record does not contain the source document identifier in the delete record.

14. The method of claim 1, wherein the collection of linked documents reside on a plurality of computers interconnected by the Internet.

15. The method of claim 1, wherein the collection of linked documents includes a first document and a second document,

wherein a document address of the first document contains information about a first host on an intranet;

wherein a document address of the second document contains information about a second host on an intranet; and

wherein the first host and the second host are distinct computer systems connected to one another.

16. A system for processing information about documents in a collection of linked documents, the system comprising:

a link log, the link log comprising a plurality of link records, wherein each link record comprises a respective source document identifier corresponding to a respective source document address and a respective list of target document identifiers corresponding to respective target document addresses; and

a global state manager configured to access the link log;

wherein the global state manager is configured to output a sorted anchor map, the sorted anchor map comprising a plurality of anchor records, each anchor record comprising a respective target document identifier and a respective list of source document identifiers; and

wherein the plurality of anchor records are ordered in the sorted anchor map based, at least in part, on their respective target document identifiers; and

wherein, for at least one anchor record, a document located at a source document address corresponding to a source document identifier in the list of source document identifiers contains at least one outbound link, the at least one outbound link pointing to a corresponding target document address, the target document address corresponding to the respective target document identifier for the at least one anchor record.

17. The system of claim 16, wherein each anchor record in the anchor map further comprises a respective list of annotations.

18. The system of claim 17, wherein the at least one outbound link is annotated by an annotation, the annotation included in the respective list of annotations for the at least one anchor record.

19. The system of claim 17, further including an indexer, wherein the indexer is configured to build an index of the collection of documents based, at least in part, on the sorted anchor map.

20. The system of claim 17, wherein at least one entry in the respective list of annotations for an anchor record in the anchor map includes a text passage and a list of attributes of the text passage.

21. The system of claim 20, wherein the text passage is determined from text within a predetermined distance of an anchor tag in a source document in the collection of documents.

22. The system of claim 17, further including a layered set of sorted anchor maps, wherein each sorted map in the layered set of anchor maps is associated with a production time.

23. The system of claim 22, further including:

a merged anchor map, the merged anchor map produced by the global state manager;

wherein the merged anchor map includes a plurality of merged anchor map records, each merged anchor record corresponding to at least one anchor record from the subset of sorted anchor maps, wherein the merged anchor records are ordered in the merged anchor map based on their respective target document identifiers.

24. The system of claim 22, wherein at least one record in at least one sorted anchor map in the layered set of anchor maps contains a delete entry.

25. The system of claim 16,

further including a sorted link map, the sorted link map comprising a plurality of link map records, each link map record comprising a respective source document identifier and a respective list of target document identifiers, wherein the link map records are ordered in the sorted link map based on respective source document identifiers;

wherein the global state manager is configure to produce the sorted link map.

26. The system of claim 16, further comprising:

a page ranker configured to access the sorted link map;

wherein the page ranker determines, based at least in part on the sorted link map, a query-independent relevance metric for a document at a source document address corresponding to a source document identifier from the sorted link map.

27. The system of claim 26, wherein the query-independent relevance metric is a page rank.

28. The system of claim 25, further including

a layered set of sorted link maps, wherein each sorted map in the layered set of link maps is associated with a production time; and

a merged link map, the merged link map produced by the global state manager;

wherein the merged link map includes a plurality of merged link map records, each merged link record corresponding to at least one link map record from a subset of the layered set of sorted link maps, wherein the merged link records are ordered in the merged link map based on respective source document identifiers.

29. A computer program product for use in conjunction with a computer system, the computer program product comprising a computer readable storage medium and a computer program mechanism therein, the computer program mechanism comprising:

a link log data structure, the link log comprising a plurality of link records, wherein each link record comprises a respective source document identifier corresponding to a respective source document address and a respective list of target document identifiers corresponding to respective target document addresses;

a global state manager module including instructions for accessing the link log; and

a sorted anchor map data structure,

wherein the global state manager module contains instructions for writing to the sorted anchor map data structure,

wherein the plurality of anchor records are ordered in the sorted anchor map data structure based, at least in part, on respective target document identifiers;

wherein, for at least one anchor record, a document located at a source document address corresponding to a source document identifier in the list of source document identifiers contains at least one outbound link, the at least one outbound link pointing to a corresponding target document address, the target document address corresponding to the respective target document identifier for the at least one anchor record.

30. The computer program product of claim 29, wherein each anchor record in the anchor map data structure further comprises a respective list of annotations.

31. The computer program product of claim 30, wherein the at least one outbound link is annotated by an annotation, the annotation included in the respective list of annotations for the at least on anchor record.

32. The computer program product of claim 30, further including an indexer module, wherein the indexer module includes instructions for building an index of the collection of documents based, at least in part, on the contents of the sorted anchor map data structure.

33. The computer program product of claim 30, wherein at least one entry in the respective list of annotations for an anchor record in the anchor map data structure includes a text passage and a list of attributes of the text passage.

34. The computer program product of claim 33, wherein the text passage is determined from text within a predetermined distance of an anchor tag in a source document in the collection of documents.

35. The computer program product of claim 30, further including a layered set of sorted anchor map data structures, wherein each sorted map in the layered set of anchor map data structures stores a respective production time.

36. The computer program product of claim 35, further including:

a merged anchor map data structure;

wherein the global state manager module includes instructions for writing to the merged anchor map data structure; and

wherein the merged anchor map data structure includes a plurality of merged anchor map records, each merged anchor record corresponding to at least one anchor record from the layered set of sorted anchor map data structures, wherein the merged anchor records are ordered in the merged anchor map based on respective target document identifiers.

37. The computer program product of claim 35, wherein at least one record in at least one sorted anchor map data structure in the layered set of anchor map data structure contains a delete entry.

38. The computer program product of claim 30,

further including a sorted link map data structure, the sorted link map data structure comprising a plurality of link map records, each link map record comprising a respective source document identifier and a respective list of target document identifiers, wherein the

link map records are ordered in the sorted link map data structure based on their respective source document identifiers;

wherein the global state manager module includes instructions for writing to the sorted link map data structure.

39. The computer program product of claim 30, further comprising:

a page ranker module;

wherein the page ranker module includes instructions for accessing the sorted link map; and

wherein the page ranker module further includes instructions for determining, based at least in part on the sorted link map, a query-independent relevance metric for a document corresponding to a source document identifier from the sorted link map data structure.

40. The computer program product of claim 39, wherein the query-independent relevance metric is a page rank.

41. The computer program product of claim 38, further including

a layered set of sorted link map data structures, wherein each sorted map data structure in the layered set of link map data structures stores a respective production time; and

a merged link map data structure, wherein the global state manager module includes instructions for writing to the merged link map data structure;

wherein the merged link map data structure includes a plurality of merged link map records, each merged link record corresponding to at least one anchor record from a subset of the layered set of sorted link map data structures, wherein the merged link records are ordered in the merged link map data structure based on respective source document identifiers.